

PATENT APPLICATION
PROCESSOR ARCHITECTURE FOR EXECUTING
TWO DIFFERENT FIXED-LENGTH INSTRUCTION SETS

Inventor(s):

Sivaram Krishnan, a citizen of the United States residing at 1723
Westbrook Avenue, Los Altos, California 94024

Mark Debbage, a citizen of the United Kingdom residing at 5636
Stevens Creek Blvd., Suite 371, Cupertino, California 95014

Sebastian Havluj Ziesler, a citizen of the United States residing at
5315 Joseph Lane, San Jose, California 95118

Kanad Roy, a citizen of the Canada residing at 610 Park View Drive,
No. 111, Santa Clara, California 95054

Andrew Craig Sturges, a citizen of the United Kingdom residing at 3
Pump Lane, Bathford, Bath BA1 7RT, UK

Kevin Ray Iadonato, a citizen of the United States residing at 678
Galleon Court, San Jose, California 95133

Prasenjit Biswas, a citizen of India residing at 20167 Pampas Court,
Saratoga, California 95070

Assignee: Hitachi Ltd.
6 Kanda Surugadai 4-chome, Chiyoda,
Tokyo, 101-8010 Japan

Entity: Large

PROCESSOR ARCHITECTURE FOR EXECUTING TWO DIFFERENT FIXED-LENGTH INSTRUCTION SETS

BACKGROUND OF THE INVENTION

5

The invention relates generally to microprocessor/microcontroller architecture, and particularly to an architecture structured to execute a first fixed-length instruction set with backward compatibility to a second, smaller fixed instruction.

10

Recent advances in the field of miniaturization and packaging in the electronics industry has provided the opportunity for the design of a variety of "embedded" products. Embedded products are typically small and hand-held, and are constructed to include micro-controllers or microprocessors for control functions. Examples of embedded products include such handheld business, consumer, and industrial devices as cell phones, pagers and personal digital assistants (PDAs).

15

20

A successful embedded design or architecture must take into consideration certain requirements such as the size and power consumption of the part to be embedded. For this reason, some micro-controllers and microprocessors for embedded products are designed to incorporate Reduced Instruction Set Computing (RISC) architecture which focuses on rapid and efficient processing of a relatively small set of instructions. Earlier RISC designs, however, used 32-bit, fixed-length instruction sets. To further minimize the processing element, designs using small fixed size, such as 16-bit were developed, enabling use of compact code to reduce the size of the instruction memory. RISC architecture coupled with small, compact code permits the design of embedded products to be simpler, smaller, and power conscious. An example of such a 16-bit architecture is disclosed in U.S. Pat. No. 5,682,545.

25

30

However, the need for more computing capability and flexibility than can be provided by a 16-bit instruction set exists, and grows, particularly when the capability for graphics is desired. To meet this need, 32-bit instruction set architectures are being made available. With such 32-bit instruction set

architectures, however, larger memory size for storing the larger 32-bit instructions is required. Larger memory size, in turn, brings with it the need for higher power consumption and more space, requirements that run counter to the design of successful embedded products.

5 Also, present 32-bit instruction set architectures provide little, if any, backward compatibility to earlier-developed, 16-bit code. As a result, substantial software investments are lost. Thus, applications using the prior, smaller, code must be either discarded or recompiled to the 32-bit instruction.

10 Thus, it can be seen that there is a need to provide a 32-bit instruction architecture that imposes a negligible impact on size and power consumption restraints, as well as providing a backward compatibility to earlier instruction set architectures.

SUMMARY OF THE INVENTION

15 Broadly, the present invention is directed to a processor element, such as a microprocessor or a micro-controller, structured to execute either a larger fixed-length instruction set architecture or an earlier-designed, smaller fixed-length instruction set architecture, thereby providing backward compatibility
20 to the smaller instruction set. Execution of the smaller instruction set is accomplished, in major part, by emulating each smaller instruction with a sequence of one or more of the larger instructions. In addition, resources (e.g., registers, status bits, and other state) of the smaller instruction set architecture are mapped to the resources of the larger instruction set environment.

25 In an embodiment of the invention, the larger instruction set architecture uses 32-bit fixed-length instructions, and the smaller instruction set uses 16-bit fixed length instructions. However, as those skilled in this art will see, the two different instruction sets may be of any length. A first group of the 16-bit instructions will each be emulated by a single 32-bit instruction sequence. A
30 second group of the 16-bit instructions are each emulated by sequences of two or more of the 32-bit instructions. Switching between the modes of execution is accomplished by branch instructions using target addresses having a bit position

(in the preferred embodiment the least significant bit (LSB)) set to a predetermined state to identify that the target of the branch is a member of one instruction set (e.g., 16-bit), or to the opposite state to identify the target as being a member of the other instruction set (32-bit).

5 The particular 16-bit instruction set architecture includes what is called a "delay slot" for branch instructions. A delay slot is the instruction immediately following a branch instruction, and is executed (if the branch instruction so indicates) while certain aspects of the branch instruction are set up, and before the branch is taken. In this manner, the penalty for the branch is
10 diminished. Emulating a 16-bit branch instruction that is accompanied by a delay slot instruction is accomplished by using a prepare to branch (PT) instruction in advance of the branch instruction that loads a target register. The branch instruction then uses the content of the target register for the branch. However, when emulating a 16-bit branch instruction with a delay slot requirement, the
15 branch is executed, but the target instruction (if the branch is taken) is held in abeyance until emulation and execution of the 16-bit delay slot instruction completes.

 The 32-bit PT instruction forms a part of a control flow mechanism that operates to provide low-penalty branching in the 32-bit instruction set
20 environment by separating notification of the processor element of the branch target from the branch instruction. This allows the processor hardware to be made aware of the branch many cycles in advance, allowing a smooth transition from the current instruction sequence to the target sequence. In addition, it obviates the need for the delay slot technique use in the 16-bit instruction set
25 architecture for minimizing branch penalties.

 A feature of the invention provides a number of general purpose registers, each 64-bits in length, for use by either the 16-bit instructions or the 32-bit instructions. However, when a general purpose register is written or loaded by a 16-bit instruction, only the low order 32-bits are used. In addition, an automatic
30 extension of the sign bit is performed when most 16-bit instructions load a general purpose register; that is, the most significant bit of the 32-bit quantity placed in the low-order bit positions of a 64-bit general purpose register are

copied to all 32 of the high-order bits of the register. The 32-bit instruction set architecture includes instructions structured to use this protocol, providing compatibility between the 16-bit and 32-bit environments.

Also, a 64-bit status register is provided for both the 16-bit
5 instruction set and the 32-bit instruction set. Predetermined bit positions of the status register are reserved for state that is mapped from the 16-bit instruction set. Other of the 16-bit state is mapped to predetermined bit positions of certain of the general purpose registers. This mapping of the 16-bit instruction set state allows the separate environment (16-bit, 32-bit) to save all necessary context on
10 task switching, and facilitates emulation of the 16-bit instructions with 32-bit instructions.

A number of advantages are achieved by the present invention. The ability to execute both 16-bit code and 32-bit code allows a processor to use the compact, 16-bit code for the mundane tasks. This, in turn, allows a saving of
15 both memory space and the other advantages attendant with that saving (e.g., smaller memory, reduced power consumption, and the like). The 32-bit code can be used when more involved tasks are needed.

Further, the ability to execute an earlier-designed 16-bit instruction set architecture provides a compatibility that permits retention of the investment
20 made in that earlier design.

The PT instruction, by providing advance notice of a branch, allows for more flexibility in the performance of branch instructions.

These and other advantages and features of the present invention will become apparent to those skilled in this art upon a reading of the following
25 detailed description which should be taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

30 Fig. 1 is a block diagram broadly illustrating a processing system employing a processor element constructed to implement the present invention;

Fig. 2 is a block diagram illustration of the instruction fetch unit (IFU) of the processor element shown in Fig. 1;

Fig. 3 is a layout of a status register contained of the branch unit shown in Fig. 2;

5 Fig. 4 is a block diagram illustration of the decoder (DEC) shown in Fig. 2;

Fig. 5 illustrates state mappings from one instruction set architecture to a second instruction set architecture; and

10 Fig 6 is a flow diagram illustrating aspects of the invention to control instruction flow.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

15 The present invention preferably provides backward compatibility to a previously-developed 16-bit fixed-length instruction set architecture. A more complete description of that architecture may be found in "SH7750 Programming Manual" (Rev. 2.0, Copyright March 4, 1999), available from Hitachi Semiconductor (America) Inc., 179 East Tasman Drive, San Jose, CA 95134.

20 Turning now to the Figures, and for the moment specifically to Fig. 1, there is illustrated, in broad form, a block diagram of the processor element (e.g., microcomputer) constructed in accordance with the teachings of the present invention. As shown in Fig. 1, a processor system, identified generally with the reference numeral 10, includes a processor element 12, an external interface 14, and a direct memory access (DMA) unit 14 interconnected by a system bus 20. Preferably, the external interface 14 is structured to connect to external memory and may also provide the processor element 12 with
25 communicative access to other processing elements (e.g., peripheral devices, communication ports, and the like).

Fig. 1 also illustrates the logical partitioning of the processor
30 element 12, showing it as including a bus interface unit (BIU) 24 that interfaces the processor unit 12 with the external interface 14 and DMA 16. The BIU 24, which handles all requests to and from a system bus 20 and an external memory (not shown) via the external interface 14, communicatively connects to an instruction flow unit (IFU) 26. The IFU 26 operates to decode instructions it

fetches from the instruction cache unit (ICU) 27, and serves as the front end to an instruction decode and execution pipeline. As will be seen, the IFU 26 contains the translation logic for emulating a 16-bit instruction set with sequences of 32-bit instructions set according to the present invention. (Hereinafter, the 16-bit instruction set architecture will be referred to as "Mode B," and the 32-bit instruction set architecture will be referred to as "Mode A.")

The BIU 24 also connects to a load-store unit (LSU) 28 of the processor element 12 which handles all memory instructions and controls operation of the data cache unit (DCU) 30. An integer/multimedia unit (IMU) 32 is included in the processor element 12 to handle all integer and multimedia instructions and forms the main datapath for the processor element 12.

In major part, the IFU 26 functions as the sequencer of the processor element 12. Its main function is to fetch instructions from the ICU 27, decode them, read operands from a register file 50 (Fig. 2), send the decoded instructions and the operands to the execution units (the IMU 32 and LSU 28), collect the results from the execution units, and write them back to the register file. Additionally, the IFU 26 issues memory requests to the BIU 24 on instruction cache misses to fill the instruction cache with the missing instructions from external memory (not shown).

Another major task of the IFU is to implement the emulation of Mode B instructions. Specifically, all Mode B instructions are translated so that the particular Mode B instruction is emulated by either one of the Mode A instructions, or a sequence of Mode A instructions. The Mode A instructions are then executed with very little change to the original Mode A instruction semantics. This approach allows the circuitry and logic necessary for implementing Mode B instruction to be isolated within a few functional logic blocks. This, in turn, has the advantage of permitting changes in the Mode B instruction set at some future date, or perhaps more importantly, being able to remove the Mode B instruction set altogether.

Fig. 2 is a block diagram of the IFU 26 illustrating it in somewhat greater detail. Because of the sequencing role played by the IFU 26 within the processor element 12, the IFU interfaces with almost every other unit of the processor element 12. The interface between the IFU 26 and both the BIU 24 and ICU 27 is established by the ICACHE) instruction cache (control (ICC) 40

which handles the loading of instructions into the ICU 27, and the flow of instructions from the ICU 27 for execution. The interface between the ICU 27 and the LSU 28 and IMU 32 provides the paths for sending/receiving instructions, operands, results, as well as all the control signals to enable the execution of instructions. In addition to these interfaces, the IFU 26 also receives external interrupt signals from an external interrupt controller 41 which samples and arbitrates external interrupts. The IFU 26 will then arbitrate the external interrupts with internal exceptions, and activate the appropriate handler to take care of the asynchronous events.

As Fig. 2 shows, the ICC 40 communicates internally with a fetch unit (FE) 42 and externally with the ICU 27 to set up accesses. Normally, the FE 42 provides an instruction fetch address and a set of control signals indicating a "fetch demand" to the ICC 40. In return, the ICC 40 sends up to two word-aligned, instruction words back to the FE 42. When the ICU 27 misses, the ICC 40 will initiate a refill cycle to the BIU 24 to load the missing cache line from the external memory (not shown). The refill occurs while the FE 42 is holding the original fetch address. Alternatively, the FE 42 may provide a "prefetch request" which requires no instruction returned, or a "fetch request," which requires no refill activity when a cache miss is experienced.

Instructions fetched from the ICU 27 by the FE 42 are first deposited in a buffer area 42a in accordance with the instruction set architecture mode of the instructions (i.e., whether Mode B or Mode A). Eventually, however, the instructions will be transported into one of two instruction buffers for application to a decode (DEC) unit 44.

When the processor element 12 is executing Mode A instructions, the DEC 44 will decode the instruction and send the decoded instruction information to the FE 42, the branch unit (BR) 46, and the pipeline control (PPC) 48, and externally to the IMU 32 and the LSU 28. The information will also allow the IMU 32 and the LSU 28 to initiate data operations without further decoding the instruction. For branch instructions, the partially decoded branch information enables the BR 46 to statically predict the direction of the branches at the earliest possible time.

When Mode B instructions are executing, all instructions will go through an additional pipeline stage: the Mode B translator 44a of the DEC 44.

The Mode B translator 44a will translate each Mode B instruction into one or multiple Mode A emulating instructions. The Mode A emulating instructions are then moved to a buffer of the DEC 44 where normal Mode A instruction decoding and execution resumes. As an example, Appendix A hereto shows, for each of
5 Mode B move and arithmetic instructions, the Mode A instruction sequences used to emulate the Mode B instruction. (The Mode B instruction set comprises many more instructions, including floating point instructions, as can be seen in the SH7750 programming manual identified above. Appendix A is used only to illustrate emulation.) As those skilled in this art will recognize, the emulation
10 sequence depends upon the particular instruction set architectures.

In addition, in order to ensure compatibility when processing 32-bit data, additional Mode A instructions are included in the Mode A instruction set for emulating the Mode B (32-bit data) instructions. These additional instructions, shown in Appendix B, operate to handle 32-bit data by retrieving only the lower
15 32 bits of the source register(s) identified in the instruction. Any result of the operation will be written to the lower 32 bits of the destination register identified in the instruction, and the sign bit of the written quantity (i.e., the most significant bit) will be extended into the upper 32 bits of the destination register.

An example of the emulation of a Mode B instruction by a Mode A instruction is illustrated by the Mode B add (ADD) instruction shown in Appendix A. This is one of the Mode B instructions that is emulated by a single Mode A instruction, an add long (add.l) instruction. In the Mode B instruction set architecture, the ADD instruction will add the contents of two 16 general purpose registers Rm, Rn to one another and store the result in the general purpose
25 register Rn. (As will be seen, the 16 general purpose registers (R₀ - R₁₅) are mapped to the low-order 32 bits of the 64-bit general purpose registers (R₀ - R₁₅) 50.) Emulation of this Mode B ADD instruction uses the Mode A add long (add.l) instruction which uses only the low-order 32-bits of the general purpose registers. Add.l operates to add the content of general purpose register Rm to the content
30 of general purpose register Rn and store the result in the low-order 32 bits of the general purpose register Rn with automatic extension of the sign bit into the high-order 32 bits of the register. Thereby, the Mode B ADD instruction is emulated by the Mode A add.l instruction to perform the same task, and obtain the same 32-bit result. (Mode A instructions use the entire 64 bits of the general purpose

registers. If a value to be written to a register is less than the full 64 bits, whether written by a Mode B instruction or a Mode A instruction, the sign of that value is extended into the upper bit positions of the register - even for most unsigned operations. This allows the result of Mode B or Mode A operation to be considered as producing a 64-bit result.) For the Mode B instruction set, as described in the SH 7750 Programming Manual identified above, the added Mode A instructions are set forth in Appendix B hereto.

An example of an emulation of a Mode B instruction by a sequence of two or more Mode A instructions is shown in Appendix A by the Mode B add-with-carry (ADDC) instruction. The ADDC instruction is similar to the ADD instruction, except that the content of the registers R_m , R_n are treated as unsigned numbers, and the sum will include a carry produced by a prior addition - stored in a 1-bit T register of the Mode B instruction set architecture. If the ADDC produces a carry, it is stored in the 1-bit T register in the Mode B environment for use by a subsequent ADDC instruction, or for other operations. This requires emulation by a sequence of Mode A instructions. (References to registers are the 64-bit general purpose registers contained in the register file 50 of Fig. 2) As can be seen, the ADDC instruction is emulated by a sequence of six Mode A instructions:

1. A Mode A add unsigned long (addz.l) instruction adds the low-32 bits of the general purpose register R_m to the general purpose register R_{63} , (which is a constant "0") and returns the result to a general purpose register R_{32} , used as a scratch pad register, with zeros extended into the high-order 32 bits of the result.

2. Next, an addz.l adds the low-32 bits of the general purpose register R_n to the general purpose register R_{63} , and returns the result to the general purpose register R_n with zeros written into the high-order 32 bits of the result.

3. Then, the Mode A add instruction adds the contents of R_n and R_{32} to one another (both of which have a 32-bit quantity in the 32 low-order bit positions, and zeros in the 32 high-order bit positions), storing the result in the register R_n .

4. The Mode A add instruction adds the result, now in register R_n to whatever carry was produced earlier and placed in the LSB of general purpose register R_{25} , and returns the result to register R_n .

5

Since the result of step 4 may have produced a carry that would have been set in the 1-bit T register, in the Mode B environment the register to which the T register is mapped (the LSB of general purpose register R_{25}) is loaded with carry during the remaining steps of the emulation:

10

5. The value held in R_n is shifted right 32 bit positions to move any carry produced from the addition into the LSB of the value and writes the result to R_{25} .

15

6. Finally, since the content of the register R_n is not a sign-extended value, the Mode A add immediate instruction adds the content to zero and returns the sign-extended result to R_n .

There are also Mode B instructions that are emulated in a single Mode A instruction or a sequence of two Mode A instructions, depending upon the registered values used by the instruction. An example of this dual personality emulation are the three Move data instructions in which the source operand is memory (MOV.B, MOV.W, and MOV.L where the source is $@R_m$). In the Mode B environment these instructions will retrieve the data in memory at the memory location specified by the content of the general purpose register R_m , add it to the content of the register R_n and return the result to the register R_n . Then, if m is not equal to n (i.e., the data is being moved to any other register than the one that held the memory address), the content of the register R_m is incremented. As can be seen in Appendix A, only one instruction is used if the data is moved from memory to the general purpose register holding the memory address of that data. If, on the other hand, the data is being moved elsewhere, the memory address is incremented by the second instruction.

30

Returning to Fig. 2, the BR 46 handles all branch related instructions. It receives the decoded branch instructions from the DEC 44,

determines whether branch conditions and target addresses are known, and proceeds to resolve/predict the branch. If the branch condition is unknown, the BR 46 will predict the branch condition statically. The predicted instruction will then be fetched and decoded. In some instances, the predicted instruction may
5 be fetched and decoded before the branch condition is resolved. When this happens, the predicted instruction will be held in the decode stage until the BR 46 is sure that the prediction is correct.

The BR 46 includes 8 target address registers 46a as well as a number of control registers, including status register (SR) 46b (Fig. 3). Branches
10 are taken in part based upon the content of one or another of the target address registers 46a. A specific target address register can be written with a target address at any time in advance of an upcoming branch instruction in preparation of the branch, using a prepare to branch (PT) instruction. As will be discussed more fully, use of the target address registers 46a to prepare for a branch in
15 advance reduces the penalty of that branch.

The (SR) 46b is a control register that contains fields to control the behavior of instructions executed by the current thread of execution. Referring for the moment to Fig. 3, the layout of SR 46b is shown. The "r" fields (bit positions 0, 2-3, 10-11, 24-25, 29, and 32-63) indicate reserved bits. Briefly, the
20 fields of SR 46b pertinent to the present invention behave as follows:

- The 1-bit fields S, Q, and M (bit positions 1, 8, and 9, respectively) are used during the emulation of Mode B instructions with Mode A instructions during certain arithmetic operations not relevant to the understanding of the present invention. These bit positions are state
25 mapped from the Mode B instruction set architecture environment for use in emulating Mode B instructions with the Mode A instruction set architecture.

- The 1-bit fields FR, SZ and PR (bit positions 14, 13, and 12, respectively) are used to provide additional operation code
30 qualification of Mode B floating-point instructions.

The Mode B instruction set architecture also uses a 1-bit T register for, among other things, keeping a carry bit resulting from unsigned add operations. The Mode B T register is, as indicated above, mapped to the LSB of the general purpose register R25. Other mappings

will be described below. It will be appreciated, however, by those skilled in this art that the particular mappings depend upon the particular instruction set architecture being emulated and the instruction set architecture performing the emulation.

5 Once instructions are decoded by the DEC 44, the PPC 48 monitors their execution through the remaining pipe stages - such as the LSU 28 and/or IMU 32. The main function of the PPC 48 is to ensure that instructions are executed smoothly and correctly and that (1) instructions will be held in the decode stage until all the source operands are ready or can be ready when
10 needed (for IMU 32 multiply-accumulate internal forwarding), (2) that all synchronization and serialization requirements imposed by the instruction as well as all internal/external events are observed, and (3) that all data operands/temporary results are forwarded correctly.

 To simplify the control logic of the PPC 48, several observations
15 and assumptions on the Mode A instruction set execution are made. One of those assumptions is that none of the IMU instructions can cause exception and all flow through the pipe stages deterministically. This assumption allows the PPC 48 to view the IMU 32 as a complex data operation engine that doesn't need to know where the input operands are coming from and where the output results
20 are going.

 Another major function of the PPC 48 is to handle non-sequential events such as instruction exceptions, external interrupts, resets, and the like. Under normal execution conditions, this part of the PPC 48 is always in the idle state. It awakens when an event occurs. The PPC 48 receives the external
25 interrupt/reset signals from an external interrupt controller (not shown), and internal exceptions from many parts of the processor element 12. In either case, the PPC 48 will clean up the pipeline, and inform the BR 46 to save core state and branches to the appropriate handler. When multiple exceptions and interrupts occur simultaneously, an exception interrupt arbitration logic 48a of the
30 PPC 48 arbitrates between them according to the architecturally defined priority.

 The general purpose registers mentioned above, including registers R_0 - R_{63} , are found in a register file (OF) 50 of the IFU 26. Each of the general purpose registers is 64-bits wide. Control of the OF 50 is by the PPC 48. Also, the general purpose register R_{63} is a 64-bit constant (a "0").

The Mode B translator 44a of the DEC 44 is responsible for translating Mode B instructions into sequences of Mode A instructions which are then conveyed to the Mode A decoder 44b of the DEC for decoding. For Mode B translation, the DEC looks at the bottom 16 bits of the instruction buffer 42a of the FE 42, and issues one Mode A instruction per cycle to emulate the Mode B instruction. The Mode A instruction is routed back to a multiplexer 43 of the FE 42 and then to the Mode A decoder 44b. A translation state is maintained within the DEC 44 to control the generation of the Mode B emulating sequences. When all emulating instructions are generated, the DEC 44 informs the FE 42 to shift to the next Mode B instruction, which can be in the top 16 bits of the instruction buffer 42a or the bottom 16 bits of the buffer.

Fig. 4 illustrates the FE42 and the DEC 44 in greater detail. As Fig. 4 shows, the instruction buffer (IB) 42a receives instructions fetched from the ICC. Instructions are pulled from the IB 42a and applied to the Mode B translator 44a of the DEC 44 and the Mode A pre-decoder 44c, depending upon the mode of operation (i.e., whether Mode B instructions are being used, and emulated by Mode A instructions, or whether only Mode A instructions are being used). Pre-decoded Mode A instructions (if operating in Mode A) or Mode A instructions from the Mode B translator (if operating in Mode B) are selected by the multiplexer 43 for application to the Mode A decoder 44b. The Mode A pre-decoder will produce the 32-bit instruction, plus some pre-decode signals. The Mode B translator will also produce Mode A 32-bit instructions plus decode signals emulating the pre-decode signals that would be produced by the Mode A pre-decoder if the Mode A instruction had been applied to it.

The FE 42 includes a Mode latch 42b that is set to indicate what mode of execution is present; i.e., are Mode A instructions being executed, or are Mode B instructions being translated to Mode A instructions for execution. The Mode latch 42b controls the multiplexer 43. As will be seen, according to the present invention the mode of instruction execution is determined by the least significant bit (LSB) of the target address of branch instructions. When operating in the Mode A environment, a switch to Mode B is performed using a Mode A unconditional branch instruction (BLINK), with the LSB of the address of the target instruction set to a "0". Switches from Mode B to Mode A are initiated by

several of the Mode B branch instructions, using a target address with an LSB set to a "1".

5 A "delay slot present" (DSP) latch 42c in the FE 42. The DSP 42c is set by a signal from the Mode B translator 44a of the DEC 44 to indicate that a Mode B branch instruction being translated is followed by a delay slot instruction that must be translated, emulated, and executed before the branch can be taken. The DSP 42e will be reset by the FE 42 when the delay slot instruction is sent to the Mode B translator 44a for translation.

10 Fig. 4 shows the DEC 44 as including the Mode B translator 44a, the Mode A decoder 44b, and a Mode A pre-decoder 44c. Mode B instructions are issued from the FE 42, buffered, and applied to the Mode B translator 44a, a state machine implemented circuit that produces, for each Mode B instruction, one or more Mode A instructions. The Mode A instructions produced by the translator 44a are passed through the multiplexor circuit 43 and, after buffering,
15 applied to the Mode A decoder 44b. The decoded instruction, (i.e., operands, instruction signals, etc.) are then conveyed to the execution units.

The operational performance of a processor element is highly dependent on the efficiency of branches. The control flow mechanism has therefore been designed to support low-penalty branching. This is achieved by
20 the present invention by separating a prepare-target (PT) instruction that notifies the CPU of the branch target from the branch instruction that causes control to flow, perhaps conditionally, to that branch target. This technique allows the hardware to be informed of branch targets many cycles in advance, allowing the hardware to prepare for a smooth transition from the current sequence of
25 instructions to the target sequence, should the branch be taken. The arrangement also allows for more flexibility in the branch instructions, since the branches now have sufficient space to encode a comprehensive set of compare operations. These are called folded-compare branches, since they contain both a compare and a branch operation in a single instruction.

30 Registers used in the Mode B instruction set architecture are typically 32-bits wide, and may be less in number (e.g., 16) than those used for the Mode A instruction set architecture (which number 64, each 64 bits wide). Thus, general purpose registers for Mode B instruction execution are mapped to the low-order 32 bits of 16 of the Mode A general purpose registers of the OF 50.

In addition, as mentioned above, signed extension is used; that is, when an operand or other expression of a Mode B instruction is written to a general purpose register of the OF 50, it is written to the lower order bits, (bit positions 0-31) with the most significant bit (bit position 31) copied in the upper bit positions (32-63). In addition, status register states used in the Mode B instruction set are mapped to specific register bits of the Mode A architecture.

An example of the mapping is illustrated in Fig. 5, which shows the state of the earlier developed Mode B instruction set architecture and the Mode A architecture state upon which it is mapped. As with the particular instruction sets, those skilled in this art will recognize that the mappings depend upon the resources available. Thus, the particular mappings shown here are exemplary only, depending upon, as they do, the instruction set architectures involved. Fig. 5 shows the mapping of Mode B state (left-most column) to Mode A state (right-most column). For example, the program counter state of the Mode B architecture is mapped to the low-order bit positions of the program counter of the Mode A architecture.

In addition to register mapping, such state as various flags are also mapped. As Fig. 5 shows, 1-bit flags are mapped to specific bit positions of one or another of the general registers of the Mode A architecture. Thus, for example, the Mode A T, S, M, and Q state/flags are respectively mapped to general purpose registers R₂₅ (bit position 0), and the SR 46b (fields S, M, and Q).

Mode A instructions, being 32-bits wide, are stored on 4-byte boundaries; and the Mode B instructions are stored on either 4-byte or 2-byte boundaries. Thus, at least two bits (the LSB and LSB + 1) are unused for addressing, and available for identifying the mode of operation. Switching between Mode A and Mode B instruction execution is accomplished using branch instructions that detect the two LSBs of the target address of the branch. When executing Mode A instructions, only an unconditional branch address (BLINK) is able to switch from the Mode A operation to Mode B operation. Thus, the mode of operation can be changed using the LSB of the target address of jump instructions used in Modes A and B instruction set architectures. A "0" in this bit position indicates Mode B target instruction, while a "1" indicates Mode A target

instruction. The LSB is used only for mode indication and does not affect the actual target address.

The earlier Mode B instruction set architecture utilized a delay slot mechanism to reduce the penalty incurred for branch operations. The delay slot is the instruction that immediately follows a branch instruction, and is executed before the branch can cause (or not cause) a transition in program flow. As indicated above, a smoother transition can be made by the PT instruction to load a target address register with the target address of a branch well ahead of the branch. However, emulation of a Mode B branch instruction with a delay slot must account for the delay slot. Accordingly when a Mode B branch instruction with a delay slot is encountered, the Mode A code sequence will take the branch, but the target instruction will not be executed until the Mode B instruction following the branch instruction (i.e., the delay slot instruction) is emulated and completed.

Fig. 6 illustrates use of aspects of the invention, including employment of the PT instruction in both the Mode A and Mode B environments, respectively, and switching from a Mode A thread 58 to a more compact Mode B thread (steps 64-84), and a return to the Mode A thread.

An understanding of the present invention may best be realized from a description of the operation of branch instructions.

Mode A to Mode A Branch:

Referring to Figs. 4 and 6, assume that while a Mode A instruction stream 58 (Fig. 6) is executing, an unconditional branch is to be taken to another Mode A instruction stream (i.e., no mode switch), using the BLINK instruction. At some time before the BLINK instruction is pulled from the IB 42a for decoding, a PT instruction will load one of the 8 target address registers 46a with the target address to which the branch is to be taken (step 60). Later, the BLINK instruction will reach the top of the IB 42a and will be sent to the Mode A pre-decoder 44c for partial decoding, and then, via the multiplexer 43, to the Mode A decoder 44b of the DEC 44 (step 62). The DEC 44 will send decoded information, including an identification of the target address register 46a containing the address of the target instruction, to the BR 46.

AT step 66, the BR 46 will read the target address from the identified target address register 46a and send it to the FE 42 with a branch command signal. Subsequently, the BR will invalidate all instructions that may be in the execution pipeline following the branch instruction.

5 Meanwhile, the FE 42 will, in step 68, issue a fetch request to the ICC 40, using the target address received from the BR 46, to fetch the target instruction from the ICU 27 (Fig. 2). The FE 42, at step 70, will check the LSB of the target address. If the LSB is a "0", the FE will know that the target instruction is a Mode B instruction. Here, however, since the target instruction is a Mode A instruction, the LSB will be a "1", and no mode change takes place. At about the same time, the contents of the IB 42a are invalidated in preparation for receipt of the instruction stream of the target instruction and the instructions that follow it. When the target instruction is received from the ICC 40, it is placed in the IB 42a, and from there sent to the DEC 44 for decoding and operation continues in step
10 72.
15

Mode Switch: Mode A to Mode B Branch:

Assume now that in a Mode A instruction sequence, a switch is to be made to the more compact code of a Mode B sequence. Here is when use of the LSB of a target address comes into play. Initially, the steps 60-68 will be the same as described above, except that step 60 sees the PT instruction loading a target address register 46a with a target address having an LSB set to a "0" to indicate that the target instruction is a Mode B instruction. Then, the BLINK branch instruction that will be used for the switch from Mode A execution to Mode B execution will be sent to the DEC 44 and decoded (step 62). After decoding the BLINK instruction, DEC 44 will send to the BR 46 the identification of the target address register 46a to use for the branch. The BR, in turn, will read the content of the identified target address register 46a, send it to the FE 42 with a branch command signal (step 66), and invalidate any instructions in the execution pipeline following the branch instruction. The FE 42 sends a fetch request, using the target address, to the ICC 40, and receives in return the target instruction(step 68). In addition, at step 70 the FE will now detect that the lower bits (i.e., the LSB) of the target address is a "0" and change its internal mode state (step 76) from Mode A to Mode B by setting the Mode latch 42b accordingly
20
25
30

to indicate Mode B operation. The output of the Mode latch 42b will control the multiplexer 43 to communicate instructions from the Mode B translator 44a to the Mode A decoder 44b.

5 The switch is now complete. The instructions will now be sent to the Mode B translator (step 78) where they are translated to the Mode A instruction(s) that will emulate the Mode B instruction.

Mode B to Mode B Branch:

10 Branches while operating in Mode B are basically as described above. The Mode B branch instruction is translated to a sequence of mode A instructions that will include a PT instruction to load a target register 46a with the address of the target instruction, followed by a Mode A branch instruction to execute the branch (e.g., a BLINK branch instruction). The exception is if the Mode B branch instruction indicates a delay slot instruction following the branch
15 instruction that must be executed before the branch can be taken. If no delay slot instruction follows the Mode B branch instruction, the steps outlined above for the Mode A branch will be performed - preceded by a PT instruction to provide the address of the target instruction.

20 If a delay slot instruction exists, however, the Mode B translator 44a will, upon decoding the branch instruction and noting that it indicates existence of a delay slot instruction, will assert a DS.d signal to the FE 42 to set a latch 42c in the FE that indicates to the FE that a delay slot is present. When the BR 46 sends the branch target address to the FE 42, the FE 42 will invalidate the all contents of the IB 42a except the delay slot instruction. The FE will request the
25 ICC 40 to fetch the target instruction, and when received place it behind the delay slot instruction - if the delay slot instruction has not yet been transferred to the DEC 44. The FE 42 will also examine the LSB of the branch target address. If it is a "0," the Mode bit 42b is left unchanged.

30 The delay slot instruction is applied to the Mode B translator and translated to produce the Mode A instruction(s) that will emulate it, then the FE 42 will reset the DSP 42c to "0." When the emulation of the delay slot instruction is complete, the branch target instruction is applied to the Mode B translator.

Mode Switch: Mode B to Mode A Branch:

Again, the initial steps taken are basically the same as set forth above, even though Mode B instructions are executing. The Mode B branch instruction will be translated by the Mode B translator to produce the Mode A instruction sequences, including a PT instruction to load a target address register with the target address (with an LSB set to a "1") of the Mode A target instruction. The Mode B translator will also issue the DS.d signal to the FE if the Mode B branch instruction has a delay slot instruction following it, setting the DSP latch 42c of the FE to indicate that a delay slot instruction exists. The BR will read the content of the target address, which will have an LSB set to a "1" to indicate that the target is a Mode A instruction, and send it to the FE 42. The BR 46 will then invalidate all instructions in the pipeline following the branch instruction, except the emulation of the delay slot instruction if it happens to be in the pipeline.

Upon receipt of the target address, the FE 42 will issue a fetch request to the ICC 40, using the target address, invalidate the content of the IB 42a, except the delay slot instruction. After the delay slot instruction is translated, the FE 42 will change its mode state by setting the Mode latch to indicate Mode A operation. All further instructions from the IB 42a, including the target instruction, will now be routed by the multiplexer 43 to the Mode A pre-decoder 44c.